DOCUMENT RESUME

ED 331 858 TM 016 379

AUTHOR Bushner, Diane E.
TITLE Fall-to-Fall Testing versus Spring-to-Spring Testing: What Is the Impact on a Local Community's Chapter 1 Evaluation?
PUB DATE Apr 91
NOTE 35p.; Paper presented at the Annual Meeting of the American Educational Research Association (Chicago, IL, April 3-7, 1991).
PUB TYPE Reports - Research/Technical (143) -- Speeches/Conference Papers (150)

EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS Comparative Analysis; *Compensatory Education; *Elementary School Students; Elementary Secondary Education; *High School Students; Parochial Schools; *Program Evaluation; Program Improvement; Public Schools; Reading Tests; *Testing Programs; Test Results
IDENTIFIERS *Education Consolidation Improvement Act Chapter 1; Testing Effects; *Testing Frequency

ABSTRACT

The impact of a decision by a local program under Chapter 1, the federally funded program of financial assistance to special educational needs of children, to test students fall-to-fall or spring-to-spring was studied. Students enrolled in a Chapter 1 reading program in 1988-89 were tested on a fall-to-spring basis, a spring-to-spring basis, and a fall-to-fall basis; and the results of the three testing points were compared. Thirteen public elementary schools, three parochial elementary schools, and one public high school were involved in Chapter 1 instruction. Spring-to-spring data were available for 319 students (54.3% of the students enrolled in the Chapter 1 program in the district studied). Data for the 1988-89 school year were available for 587 students in grades 2 through 12. The time of testing produced differences in both evaluation and program improvement data. The magnitude of the Normal Curve Equivalent gains varied; different schools were identified for program improvement based on the time of testing. A summer decline in scores was documented for all time intervals. Trends from the spring-to-spring data resembled the no-longer-allowed fall-to-spring data. The requirements of program improvement negatively impacted small and mid-sized schools. Questions were raised about the stress placed on a single normative test score for program improvement. Recommendations were made for changes in the program improvement aspects of Chapter 1. Nine tables present study data, and one figure depicts the reading profile summary for an individual student. (SLD)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED331858

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☑ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

DIANE E. BUSHNER

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

Fall-to-Fall Testing Versus Spring-to-Spring Testing: What Is the Impact on a Local Community's Chapter 1 Evaluation?

Diane E. Bushner
Cambridge Public Schools
Cambridge, Massachusetts

Paper presented at the meeting of
the American Education Research Association
Chicago, Illinois: April 1991

TM016379

BEST COPY AVAILABLE

Summary

The changes in Chapter 1 evaluation and program requirements are the focus of this paper. The impact of the decision by a local Chapter 1 Program to test fall-to-fall or spring-to-spring is studied. Students who were enrolled in a Chapter 1 reading program during the 1988-1989 school year were tested on a fall-to-spring basis, a spring-to-spring basis, and a fall-to-fall basis and the results of the three testing points are contrasted.

The data revealed that the time of testing produced differences in both evaluation and program improvement data. The magnitude of the Normal Curve Equivalent (NCE) gains varied; different schools were identified for program improvement based on the time of testing; a summer decline in scores was documented in all time intervals; trends from the spring-to-spring data resembled the no longer allowed fall-to-spring data; the requirements of program improvement negatively impacted small and mid sized schools; and questions were raised regarding the stress placed on a single normative test score for program improvement. The paper concludes with five recommendations for change in the future reauthorization of the evaluation and program improvement statue of Chapter 1.

Fall-to-Fall Testing Versus Spring-to-Spring Testing: What Is the Impact on a Local Community's Chapter 1 Evaluation?

Chapter 1, a federally funded program of financial assistance to meet the special educational needs of children, dates back to the passage of the Elementary and Secondary School Act of 1965. From 1965 to the present, Congress has reauthorized Chapter 1 periodically, and most recently under the Augustus F. Hawkins-Robert T. Stafford Elementary and Secondary School Improvement Amendments of 1988.

Amendments of 1988

Some of the changed requirements, which emanate from the Amendments of 1988, are in the area of evaluation and the new area of program improvement. These two concepts, evaluation and program impr ment, are the focus of this paper, as they effect the local implementation and day to day functioning of a Chapter 1 reading program. Essentially, Chapter 1 evaluation requirements aggregate data by grade level across each community, but the results are not binding. Chapter 1 program improvement takes the same aggregated data and analyzes it again by school across each community. The results are binding on individual schools.

The rules and regulations (Education Funding Research Council, 1990), issued by the Department of Education in May 19, 1989, expand upon the statue.

Subpart H (34 CFR 200.80) lists the National Evaluation Standards for use by a Local Education Authority (LEA). The standards include the requirements for evaluating Chapter 1 participants in grades 2 through 12; for assessing reading achievement in both basic and more advanced skills; for measuring achievement over a period of approximately twelve months, either spring-to-spring or fall-to -fall testing intervals; and for using aggregated tests with national norms or if a test is without national norms, then equating the test to a nationally normed test.

Program improvement (34 CFR 200.38) requires each community with a Chapter 1 Program to conduct an annual review for each project school on the effectiveness of the Chapter 1 project. Substantial progress, measured in aggregate performance and in other desired outcomes, determines the effectiveness of each project. The local community defines other desired outcomes in the project application. Annually, each state, through the committee of local practitioners, sets the minimum aggregate performance standard in terms of Normal Curve Equivalent (NCE) units. This standard in NCE units applies to the entire state.

In practical terms, both evaluation and program improvement require score aggregation of normative test scores or the equating of non-normative scores to a normative test score. The administration of the testing occurs at approximately twelve month intervals. The students in this study have participated in a Chapter 1 reading program during the 1988-1989 school year in

Cambridge, Massachusetts, an urban community. The Chapter 1 reading program serves students in grades 3 through 12. The school system currently has 7,713 students with schools organized into eight grades of elementary school and four grades of high school. There are thirteen elementary schools, ranging in size from 171 students to 703 students, and one comprehensive high school. Three parochial schools also participate in Chapter 1 instruction. The parochial elementary schools elect to receive Chapter 1 reading instruction in grades 1 through 6.

Local Implementation of the Testing

The Chapter 1 reading program in this study has a history of evaluating the instructional program on a fall-to-spring basis. As the focus of the Chapter 1 reading program is reading comprehension, the reading comprehension subtest has constituted the basis for program evaluation since 1979. Under the current regulations the reading comprehension subtest may measure the evaluation of both basic and advanced skills. During the 1988-1989 school year, the plan for the evaluation design was to continue to use the Metropolitan Achievement Test, sixth edition, reading comprehension subtest on a fall-to-spring basis. School year 1988-1989 was to be a change over year from one set of regulations to the other, with either set of regulations acceptable.

In April of 1989, as the spring testing was in process, a rumor spread that the fall-to-spring data was no longer acceptable. Later, this rumor was proven to

be just that. Officially, Cambridge submitted fall-to-spring data for the 1988-1989 school year (Casper, 1989). To comply with this potential change, while avoiding undo confusion in the Chapter 1 classrooms, the program decided to take another look at the Spring 1988 testing and to identify students tested in both years. The matched scores would create spring-to-spring testing with Spring 1988 testing serving as the pretest and Spring 1989 testing serving as the posttest. A discussion of the spring-to-spring testing will occur later.

With the matching of spring-to-spring scores, the missing third grade public school scores became apparent. As the third graders began service in September 1988, they were not part of the testing in the previous spring. Only third graders who attended parochial schools and public school repeaters comprised the tally of third graders with two test scores. To solve the problem of missing third grade scores, plus an intuitive belief by some Chapter 1 staff members that the fall was a better time to test; the program developed tentative plans for fall-to-fall testing for the next school year. In the fall of 1989, the testing of students occurred with the intended purpose of using this score as next year's pretest score.

Thus, the Cambridge Chapter 1 program collected test scores on essentially the same students for four data points: Spring 1988, Fall 1988, Spring 1989, and Fall 1989. It was unusual to have four test scores on students who participated in a specific year. To take full advantage of the data, further analyses occurred.

All aggregated data presented in the following tables is expressed in NCE units, which facilitated the aggregation of scores across grades and schools.

Design and Test Scores

Spring 1988-to-Spring 1989 Evaluation Testing

Spring 1988 and Spring 1989 data tested three hundred-nineteen students. During the 1988-1989 school year, there were over five hundred-eighty-seven students enrolled in the reading program. The spring-to-spring analysis tested only 54.3% of the students enrolled in the program. Initially, there was no intention to match scores between the two spring tests. Thus, there occurred no attempt to locate students who were still in the school system, but no longer enrolled in Chapter 1 for the test.

Insert Table 1 about here.

In the Spring 1988 and Spring 1989 evaluation data, grades 2 through 10 displayed positive mean gain scores in reading comprehension in terms of NCE units. Although grades 2 and 3 had large mean gain scores, there were only a small number of students tested. This warranted no further discussion. The remaining gains in NCE units ranged from 10.51 NCE units for grade 4 to 0.92 NCE units for grade 10. Grades 11 and 12 had negative mean scores with a small number of students tested. In summary, all Chapter 1 students tested between Spring 1988 and Spring 1989 showed a mean gain of 5.78 NCE units on the

Metropolitan Achievement Test, reading comprehension subtest.

The scores on the Spring 1988-to-Spring 1989 testing contained similar trends in comparison to the Fall 1988-to-Spring 1989 testing. On the Fall 1988

Insert Table 2 about here.

---

and Spring 1989 evaluation data, all grades had positive mean gain scores in reading comprehension. Again, no further discussion will occur concerning grades with small numbers of students tested. The mean gains ranged from 9.43 NCE units for grade 6 to 3.74 NCE units for grade 3. In total all students tested in the Fall 1988-to-Spring 1989 showed a mean gain of 7.24 NCE units on the Metropolitan Achievement Test, reading comprehension subtest.

Although there is almost a 2:1 ratio of students tested Fall 1988-to-Spring 1989 as compared with Spring 1988-to-Spring 1989, the trend of the aggregated data is similar. The mean NCE average for the 319 students tested in the Spring 1989 as part of the spring-to-spring testing was 34.95 NCE units, while the mean NCE average for the 587 students tested in the Spring 1989 as part of the fall-to-spring testing was 35.02 NCE units. On the Spring 1989 scores, which served as posttest scores, a larger average standard deviation score occurred than in the comparison pretest scores. With instruction, the spread of scores grew wider. The pretest scores were also similar, 29.60 NCE units in Spring 1988 and 27.78 NCE units in Fall 1988.

Program Improvement in Spring 1989

The committee of local practitioners set a no loss of NCE units as the state standard for program improvement during 1988-1989. On Fall 1988-to-Spring

---

Insert Table 3 about here.

1989 data, all schools made positive mean NCE gains between the two testing points. The only school with a gain less than three NCE units was a small school that had been grandfathered into Chapter 1 eligibility for the 1988-1989 school year. The school lost Chapter 1 eligibility for the next school year. Using the Spring 1988-to-Spring 1989 data, all target schools except one, displayed a positive NCE gain.

Insert Table 4 about here.

School EL-J tested only 9 students on the two data points because it was not a target school in the previous year. School EL-K, a primary grade school, had only one third grade repeater with the pretest and posttest scores. One non-target school with a small "follow the child" program had a negative NCE score along with a parochial school that tested only one student in the twelve month interval.

Essentially, both the fall-to-spring and spring-to-spring analyses of the individual schools for program improvement showed a strong gain in NCE units.

One problem, which hampered statistical aggregation, was the small Chapter 1 enrollment at many of the schools. This caused the reliability of some of the findings to be questionable.

Fall 1988-to-Fall 1989 Evaluation Testing

A question, raised by the absence in spring-to-spring test data, was what impact would the testing of students have who no longer attended Chapter 1. In preparing to test the Fall 1989 group, the program attempted to test all students who participated in the program during the previous fall, whether or not they continued in Chapter 1 during the new school year.

The Fall 1988-to-Fall 1989 data tested five hundred seven students. This number included students who participated in Chapter 1 during the 1988-1989 school year but who may or may not have continued in Chapter 1 during the 1989-1990 school year. The program made a deliberate attempt to locate students who were still in the school system for the Fall 1989 test.

Insert Table 5 about here.

On the Fall 1988 and Fall 1989 evaluation data, grades 3, 4, 6, 7, 9, 10, and 12 displayed positive mean gain scores in terms of NCE units. Grades 3 and 12 had small numbers of students tested and no discussion will center on these scores. The differences in scores ranged from a gain of 3.91 NCE units at Grade 6 to a loss of 3.41 NCE units at Grade 11. Grades 5 and 8 had good sample sizes,

but both grade levels scored a loss in NCE units. On the fall-to-fall testing the gains at specific grade levels were much smaller than on the spring scores and some grade levels experienced a loss in NCE units. In total, all students tested in Fall 1988-to-Fall 1989 showed a mean gain of 1.03 NCE units on the Metropolitan Achievement Test, reading comprehension subtest.

One important difference in the administration of the fall-to-fall testing was the concerted effort to identify students enrolled in Chapter 1 during the previous school year and to include these students in part of the Fall 1989 testing. As teachers reported that some of the "test only" students did not take the testing seriously, nor did they try hard, the questionable validity of these scores arose. An additional analysis removed the "test only" students who were not continuing in Chapter 1 for 1989-1990 school year from the data.

Insert Table 6 about here.

Two hundred-seventy-seven students continued in Chapter 1 into the 1989-1990 school year. This amounted to 54.6% of the populated tested in Fall 1989. Some students did not continue in Chapter 1 because they had different program placements, others had scored out of Chapter 1, others no longer lived in a target area, and at some parochial schools, others no longer attended an eligible grade for service.

For the students who remained in Chapter 1 in 1889-1990, the trend

continued toward more grade levels with a loss in NCE units. In addition to Grades 5 and 8, Grades 4 and 10 had a loss in NCE units. Grades 6, 7, and 9 continued positive NCE gains in both summaries of fall-to-fall data. The continuing Chapter 1 students who were tested in Fall 1988 and Fall 1989 showed a mean gain of 1.67 NCE units on the Metropolitan Achievement Test, reading comprehension subtest.

Although there is almost a 2:1 ratio of students tested in the total Fall 1988-to-Fall 1989 group as compared with the continuing Chapter 1 Fall 1988-to-Fall 1989 group, the trend of the aggregated data is similar. As might be expected, continuing Chapter 1 students had a lower mean NCE score on the two testing points, but the total group did not score substantially higher. The Fall 1988 mean scores for continuing Chapter 1 students were 25.54 NCE units compared with a mean of 28.45 NCE units for all students tested. Fall 1989 testing had a similar trend. The amount of change as measured in NCE units was also similar; continuing Chapter 1 students scored a 1.67 NCE unit gain while the total group scored a 1.03 NCE unit gain. The overall mean gain in NCE units on the fall-to-fall testing was considerably lower than the spring-to-spring mean gain of 5.78 NCE units.

Fall 1988-to-Fall 1989 Program Improvement

Consistent with the requirements of program improvement, the fall-to-fall data aggregation occurred by school. In the total Fall 1988-to-Fall 1989 data,

target schools EL-C and EL-D both showed a mean loss; -1.35 NCE units for

---

Insert Table 7 about here.

---

school EL-C and -1.41 NCE units for school EL-D. This was in contrast to the spring-to-spring data (refer to Table 4) where school EL-C had a gain of 6.39 NCE units and school EL-D had a gain of 4.96 NCE units. Both schools contained a typical size Chapter 1 program for a school system, which does not qualify for school wide projects.

Two other schools with rather small numbers of students tested, also had NCE losses. Specifically EL-I, a school with "follow the child" service, and EL-Q, a parochial school, both experienced a loss between the fall testing points. However, both schools, EL-I and EL-Q, experienced changes in Chapter 1 teachers between the two school years. In a comparison of Fall 1989 and Spring 1989 data, all schools showed a decline.

The same trends appeared when only the continuing Chapter 1 students are aggregated together. The absolute value of the NCE units changed. Target

---

Insert Table 8 about here.

---

schools EL-C and EL-D continued to show a mean loss of NCE units, as well as the two schools with the Chapter 1 staffing changes, EL-I and EL-Q. In addition parochial school EL-O also had a mean loss in NCE units. Many of the students

continuing in Chapter 1 in school EL-O came from the parochial school, which closed the previous June. Thus, many of these students also experienced a staffing change. Another issue, to be considered in schools EL-I, EL-Q and EL-O, is the impact of the small number of students tested.

Discussion of the Data

Due to an administrative decision, to test in the fall of the year or to test in the spring of the year, the aggregated data varied for both Chapter 1 evaluation and Chapter 1 program improvement. Essentially the data portrayed the same group of students who participated in the Chapter 1 reading program during the 1988-1989 school year. One of the reasons typically cited for no longer allowing fall-to-spring data was due to the problem of summer loss in achievement. Supposedly, the data collected within twelve month intervals would moderate this issue. Given the data presented, this is questionable. The fall-to-fall data suffered from the same summer decline phenomena and compared the low ebbs of the evaluation data. The spring-to-spring data showed evidence of learning, as the high ebbs are compared. In addition, the spring-to-spring data, although a lower mean gain score, displayed similar trends to the fall-to-spring data that no longer may be used. To vary the time interval in the aggregation of the data does not change the proficiency level.

A limitation of this study was the fact that the comparison data of fall-to-fall; fall-to-spring; and spring-to-spring testing intervals for the same school year

of Chapter 1 students were available for only one year. However, when the Chapter 1 teachers learned of the Fall 1988-to-Fall 1989 data, they decided to reverse the earlier vote and to use spring-to-spring data. The official data reported for school year 1989-1990 was spring-to-spring data, using Spring 1989 data as the pretest and Spring 1990 data as the posttest.

Insert Table 9 about here.

The Spring 1989-to-Spring 1990 data were consistent with earlier Spring 1988-to-Spring 1989 data. The aggregate change from Spring 1989-to-Spring 1990 was 3.0 NCE units and moderately above the expected one year gain. Also, individual schools exhibited positive NCE gains. Being a public school system and not primarily a research institution, fall Chapter 1 testing no longer was administered, nor continued to confirm or refute the findings from Fall 1988-to-Fall 1989.

Another issue was the importance of a particular Chapter 1 teacher to the Chapter 1 students. In schools where staffing changes occurred during the summer of 1989, these schools had a mean loss in NCE units on the fall-to-fall testing. Another factor not analyzed, was how many of the students tested in the fall were tested by a Chapter 1 teacher with whom the student did not study in the previous school year. The role or the impact of teacher bonding, specifically for disadvantaged or at-risk students, needs to be looked at systematically, especially

as it may influence both testing and instruction. How much of the decline, in the fall-to-fall scores was due to a summer decline or was due to the impact of a change of teachers, remains unknown. This question needed to be factored out to determine whether if it were summer decline. The pattern of summer decline suggests support for the twelve month school as a way to prevent the cyclic decline.

A staff development focus of the current year in the Cambridge Chapter 1 Program emphasizes the role of metacognition. Chapter 1 teachers are trying to help students become instructionally independent rather than fostering teacher dependence.

The program improvement regulations, which required the analysis of aggregated scores on a school by school basis, impacted adversely on school districts that maintain small to medium sized elementary schools. Given the limited size of the school enrollment, there is a relative small pool of students at each school who qualify for and participate in Chapter 1. When using such small numbers, the reliability of the aggregated results may be questionable.

Program improvement regulations placed heavy reliance on a single normative test score. Under program improvement regulations an individual school must show significant growth on the basis of both the aggregated normative score and other desired outcomes. From various sources (National Commission on Testing and Public Policy, 1990; Valencia, McGinley & Pearson,

1990) questions were raised about the need for multiple sources of assessment data or for new ways of looking at assessment. Yet within the program improvement regulations, Chapter 1 placed emphasis on the single normative score.

How reliable is program improvement data, when multiple test points are available? In Table 4 on Spring 1988-to-Spring 1989 data, school EL-C had a mean gain of 6.39 NCE units while school EL-D had a mean gain of 4.96 NCE units. Using the t-test for nonindependent samples (Ary, Jacobs, & Razavieh, 1972), both schools with the spring data scored a statistically significant gain at the .01 level. In contract, the analysis of the Fall 1988-to-Fall 1989 scores (refer to Table 7 and Table 8) showed both schools suffered a loss in NCE units.

These two schools represent average size schools with one and a half Chapter 1 teachers assigned to each school. By an administrative decision of fall versus spring testing, the aggregation of test scores looks entirely different in these two schools. Using spring data, schools EL-C and EL-D will be heralded as successful Chapter 1 schools with strong NCE unit gains, but using fall data the same schools will be designated for program improvement and their instructional program will be judged ineffective. The same teachers, students, and curriculum are present, but the time of data aggregation changed.

Goodlad (1990) in a journal article, questioned why our schools don't get much better. He discussed seven propositions that he had written a decade

earlier. One of the propositions stated that the norm by which the performance of schools is judged is entirely inadequate. Is this not this the problem with the Chapter 1 program improvement statue? Although there is mention of flexibility to design more meaningful measures under the rubric of "other desired outcomes," this becomes inconsequential if the aggregated normative scores do not show a sizable gain in mean NCE units. Regardless of how substantial the progress is in the other desired outcomes, the school enters into program improvement, if the aggregated normative scores do not meet the state standard in terms of NCE units. A normative standard of NCE units, by which the success of a school's Chapter 1 program is judged, is inadequate. Normative test scores drive this standard, rather than sound instructional practices.

During the 1989-1990 school year, attempts occurred to capture the priorities of the Cambridge Chapter 1 reading program and to translate these priorities into other desired outcomes. The program created an End of the Year

---

Insert Figure 1 about here.

Profile Summary Form in reading (Bushner, 1990). The form summarized the year-long reading curriculum priorities. This profile structured the review of individual students who did not show a gain in NCE units between the spring-to-spring testing. The Chapter 1 director met with each Chapter 1 teacher in late May or early June to discuss these students.

During the discussions with Chapter 1 teachers, the director recorded anecdotal comments as the teachers reviewed the individual student's Chapter 1 folder. The director aggregated the comments from the teachers on the reading comprehension and composing section across the program using the following scale. "One" meant the Chapter 1 teacher substantiated that the student was still functioning at the baseline, or the same range as the student entered the program in September; "two" meant the Chapter 1 teacher documented some improvement that the student made during the year; and "three" meant the student functioned at the average range for the grade in which the student was enrolled.

By June 1990, discussions with teachers focused on one hundred-thirteen students who did not make NCE gains from pretest to posttest. Chapter 1 teachers rated two-thirds of this group as demonstrating some improvement in reading on the basis of narrative and expository comprehension and writing. The End of the Year Profile Summary Form provided an opportunity to discuss the reading performance of Chapter 1 students from a focused curriculum perspective. The use of a rating scale allowed the results of an open ended process to be aggregated across the program. For the 1990-1991 school year, the expansion of this process will occur. Plans are in process to link the End of the Year Summary Profile to the other desired outcomes for program improvement.

Recommendations

1.) The program improvement statue of Chapter 1 needs to be written in a

positive manner rather than the present negative manner. One can always improve, yet waiting for a numeric documentation of failure to occur negatively frames the program improvement requirements. This often fosters hostility among staff. Whereas a positive statue, where schools in a district cycle through the program improvement process, ensures that all schools benefit from program improvement and this approach may foster greater collegial collaboration among staff.

2.) The specific NCE unit changes between pretest and posttest may vary due the time of the testing, may be altered by the addition or deletion of a few students in the testing pool, may be influenced by staffing changes, and other subtle factors. The statistical numbers that are computed do not adequately or reliably rank order the schools in terms of need. As consumers of educational test data, we need to refrain from giving undo power to single instances of computed numbers. The practice of using school mean change scores in NCE units to determine eligibility for program improvement should be eliminated. By relating program improvement to aggregated normative data, whatever problems are associated with this type of instrument, also are carried forth into the program improvement efforts.

3.) In school districts that do not administer a consistent normative test at every grade level, the issue of testing students, who no longer are enrolled in the Chapter 1 program, is very problematic and often produces unreliable results.

This need to test larger numbers of students who no longer are enrolled in Chapter 1 is increased by the minimum twelve month testing interval. The same phenomena of the rise and fall of test scores occurred across the twelve month interval as it did in the school year testing interval. The same trends in scores occurred across the spring-to-spring data as the fall-to-spring data. Return to the Chapter 1 regulations the option to use fall-to-spring evaluation data.

4.) Provide greater incentive to local school districts that want to devise alternative measures of other desired outcomes. The aggregation of other desired outcomes without reference to national norms should suffice for program improvement analysis on a school by school basis. Program improvement efforts should be meaningful to individual schools and the district. Encourage and support school districts to submit innovative proposals to devise alternative measures of other desired outcomes. For the purpose of maintaining the national evaluation database of Chapter 1, retain the evaluation requirements that aggregate normative scores by grade level.

5.) Other school districts are encouraged to replicate the testing of a single year of Chapter 1 students at different times: fall-to-spring; spring-to-spring; and fall-to-fall to confirm or refute the data in this study. We need to have a greater understanding of the impact that specific administrative decisions place on our students.

References

Ary, D., Jacobs, L.C. & Razavieh, A. (1972). Introduction to research in education. New York: Holt, Rinehart and Winston.

Bushner, D.E. (1990). Chapter 1 evaluation for 1989-1990. Unpublished manuscript.

Casper, C.P. (1989). Final evaluation report for the Chapter 1 reading and math program in the Cambridge Public Schools. Unpublished manuscript.

Educational Funding Research Council. (1990). Chapter 1 handbook: Understanding and implementing the Chapter 1 Program. Arlington, VA: Educational Funding Research Council.

Goodlad, J.I. (1990). Why our schools don't get much better--and how they might. Teacher Education Quarterly, 17 (4), 5-21.

National Commission of Testing and Public Policy. (1990). From gatekeeper to gateway: Transforming testing in America. Chestnut Hill, MA: Boston College.

Valencia, S.W., McGinley, W. & Pearson, P.D. (1990). Asessing reading and writing. In G.G.Duffy (Ed.) , Reading in the middle school. Newark, DE: International Reading Association.

References

Ary, D., Jacobs, L.C. & Razavieh, A. (1972). Introduction to research in education. New York: Holt, Rinehart and Winston.

Bushner, D.E. (1990). Chapter 1 evaluation for 1989-1990. Unpublished manuscript.

Casper, C.P. (1989). Final evaluation report for the Chapter 1 reading and math program in the Cambridge Public Schools. Unpublished manuscript.

Educational Funding Research Council. (1990). Chapter 1 handbook: Understanding and implementing the Chapter 1 Program. Arlington, VA: Educational Funding Research Council. (see Footnote 1)

Goodlad, J.I. (1990). Why our schools don't get much better--and how they might. Teacher Education Quarterly, 17 (4), 5-21.

National Commission of Testing and Public Policy. (1990). From gatekeeper to gateway: Transforming testing in America. Chestnut Hill, MA: Boston College.

Valencia, S.W., McGinley, W. & Pearson, P.D. (1990). Asessing reading and writing. In G.G.Duffy (Ed.) , Reading in the middle school. Newark, DE: International Reading Association.

Footnotes

1--A compilation of Chapter 1 source documents is contained in this reference. The Education Funding Research Council updates yearly the handbook and insures that the latest edition of Chapter 1 legislation, regulations, and guidelines are included. All sections relating to Chapter 1 from the Code of Federal Regulations (CFR) are part of the handbook.

TABLE 1
Spring 1989 and Spring 1988 NCE Scores
on the Metropolitan Achievement Test,
Reading Comprehension Subtest
Data by Grade

School	Number Tested	Spring 1989 Mean	Spring 1989 S.D.	Spring 1988 Mean	Spring 1988 S.D.	Difference Mean	Difference S.D.
Grade 1	0	----	----	----	----	----	----
Grade 2	5	45.24	13.64	18.86	9.10	26.38	4.90
Grade 3	11	37.02	13.95	24.38	13.84	12.64	15.16
Grade 4	58	38.28	11.85	27.77	9.36	10.51	9.88
Grade 5	58	34.71	12.59	29.56	9.42	5.15	10.82
Grade 6	52	35.79	10.59	29.82	10.93	5.97	10.93
Grade 7	45	37.18	10.82	34.26	8.86	2.92	9.06
Grade 8	31	31.58	11.75	29.23	11.22	2.35	10.90
Grade 9	36	35.60	14.63	30.29	10.65	5.31	15.53
Grade 10	16	18.93	13.49	18.01	11.11	0.92	10.20
Grade 11	6	23.88	12.28	33.07	16.90	-9.18	9.97
Grade 12	1	33.70	----	37.10	----	-3.40	----
Summary							
Program							
Total	319	34.95	12.84	29.16	10.82	5.78	11.88

TABLE 2
Spring 1989 and Fall 1988 NCE Scores
on the Metropolitan Achievement Test,
Reading Comprehension Subtest
Data by Grade

	Number	Spring 1989		Fall 1988		Difference	
Grade	Tested	Mean	S.D.	Mean	S.D.	Mean	S.D.
Grade 1	3	39.90	11.31	34.77	13.93	5.13	10.90
Grade 2	12	45.70	11.21	27.88	9.20	17.82	14.56
Grade 3	107	33.94	14.45	30.19	11.74	3.74	10.40
Grade 4	100	38.60	12.07	29.78	11.89	8.82	9.84
Grade 5	69	33.68	12.48	24.80	12.39	8.88	13.59
Grade 6	87	35.92	12.72	26.49	11.14	9.43	11.59
Grade 7	71	38.23	11.86	32.03	10.70	6.20	10.43
Grade 8	51	29.39	13.20	22.59	12.55	6.80	9.03
Grade 9	52	31.93	17.21	26.49	16.66	5.44	9.38
Grade 10	24	30.28	17.88	23.87	16.95	6.41	10.06
Grade 11	9	33.46	6.09	25.71	6.33	7.74	8.69
Grade 12	2	22.05	16.48	16.65	22.13	5.40	5.66
Summary							
Program Total	587	35.02	13.77	27.78	12.65	7.24	10.98

TABLE 3
Spring 1989 and Fall 1988 NCE Scores
on the Metropolitan Achievement Test,
Reading Comprehension Subtest
Data by School

School	Number Tested	Spring 1989 Mean	Spring 1989 S.D.	Fall 1988 Mean	Fall 1988 S.D.	Difference Mean	Difference S.D.
Target Schools							
HS-A	79	32.67	15.84	26.64	14.88	6.03	9.23
EL-C	57	43.04	12.17	37.73	12.57	5.31	10.43
EL-D	57	34.48	11.83	26.58	9.17	7.89	11.37
EL-F*	11	34.50	10.04	34.40	10.77	0.10	6.65
EL-G	135	33.87	13.65	25.95	12.15	7.93	12.22
EL-H	55	36.02	12.29	27.73	9.97	8.29	12.30
EL-J	30	25.32	11.33	19.62	12.49	5.70	8.63
EL-K	18	31.43	8.38	26.61	8.93	4.82	7.68
EL-L	26	34.23	11.78	26.01	7.49	8.22	9.87
Follow the Child Service in Non-Target Schools							
EL-E	11	27.02	13.37	29.64	15.83	7.18	7.02
EL-I	6	40.42	19.18	30.28	12.81	10.13	17.11
EL-M	12	38.08	12.20	26.94	13.99	11.14	10.10
EL-N	20	42.38	12.47	32.73	11.12	9.65	6.38
Parochial Schools							
HS-B	8	18.91	18.32	13.82	22.03	5.09	11.04
EL-O	13	41.88	10.66	30.75	10.05	11.13	17.72
EL-P	41	38.56	12.70	30.20	10.11	8.36	9.53
EL-Q	8	35.58	15.93	32.33	10.23	3.25	18.34
Program Total	587	35.02	13.77	27.78	12.65	7.24	10.98

*Grandfathered into eligibility for 1988-1989 school year.

TABLE 4
Spring 1989 and Spring 1988 NCE Scores
on the Metropolitan Achievement Test,
Reading Comprehension Subtest
Data by School

School	Number Tested	Spring 1989 Mean	Spring 1989 S.D.	Spring 1988 Mean	Spring 1988 S.D.	Difference Mean	Difference S.D.
Target Schools							
HS-A	57	29.76	15.97	27.08	12.76	2.68	14.40
EL-C	29	40.46	12.90	34.07	12.63	6.39	12.69
EL-D	31	35.34	11.20	30.38	8.30	4.96	9.60
EL-F*	3	37.80	9.54	31.30	8.66	6.50	2.65
EL-G	79	35.86	12.19	28.84	10.11	7.02	11.51
EL-H	43	34.36	12.31	25.56	8.20	8.80	9.95
EL-J	9	28.76	8.26	29.00	14.12	-0.24	12.48
EL-K	1	38.30	----	6.70	----	31.60	----
EL-L	15	34.55	10.75	33.43	8.22	1.11	9.04
Follow the Child Service in Non-Target Schools							
EL-E	6	28.05	7.07	12.63	7.41	15.42	6.62
EL-I	4	31.23	12.75	35.63	8.37	-4.40	13.39
EL-M	3	29.43	3.58	28.40	4.69	1.03	7.35
EL-N	10	40.59	5.25	38.05	4.50	2.54	5.37
Parochial Schools							
HS-B	1	31.50	----	33.70	----	-2.20	----
EL-O	6	43.42	12.31	33.08	5.75	10.33	13.44
EL-P	22	39.29	12.28	30.60	9.96	8.69	12.15
EL-Q	0	----	----	----	----	----	----
Program Total	319	34.95	12.84	29.16	10.82	5.78	11.88

*Grandfathered into eligibility for 1988-1989 school year.

Table 5
Fall 1989 and Fall 1988 NCE Scores
on the Metropolitan Achievement Test,
Reading Comprehension Subtest
All Students Tested
Data by Grade

School	Number Tested	Fall 1989 Mean	Fall 1989 S.D.	Fall 1988 Mean	Fall 1988 S.D.	Difference Mean	Difference S.D.
Grade 2	3	13.13	5.65	34.77	13.93	-21.63	13.52
Grade 3	8	38.33	8.06	31.46	7.58	6.86	4.79
Grade 4	103	30.26	13.33	29.77	11.46	0.49	11.71
Grade 5	80	29.64	13.34	30.45	10.93	-0.81	13.25
Grade 6	64	28.81	10.10	24.90	13.13	3.91	12.10
Grade 7	79	29.91	10.89	27.23	11.45	2.67	10.05
Grade 8	65	30.75	11.98	32.39	11.29	-1.64	11.57
Grade 9	41	27.52	12.28	24.00	11.92	3.51	11.73
Grade 10	41	28.49	13.02	27.91	15.90	0.57	10.11
Grade 11	15	25.69	10.29	29.11	12.60	-3.41	12.93
Grade 12	8	28.13	14.69	18.91	13.60	9.21	6.47
Summary							
Program Total	507	29.48	12.21	28.45	12.29	1.03	11.82

Table 6
Fall 1989 and Fall 1988 NCE Scores
on the Metropolitan Achievement Test,
Reading Comprehension Subtest
Continuing Chapter 1 Students
Data by Grade

School	Number Tested	Fall 1989 Mean	Fall 1989 S.D.	Fall 1988 Mean	Fall 1988 S.D.	Difference Mean	Difference S.D.
Grade 2	0	----	----	----	----	----	----
Grade 3	3	29.60	1.71	24.57	6.05	5.03	7.46
Grade 4	62	27.69	10.13	27.95	8.16	-0.27	10.06
Grade 5	47	28.27	12.48	28.90	9.68	-0.63	12.42
Grade 6	42	27.41	7.90	21.48	11.34	5.93	12.19
Grade 7	38	28.26	9.47	22.44	9.15	5.82	10.12
Grade 8	32	25.62	10.07	27.75	11.09	-2.13	10.92
Grade 9	24	28.00	13.59	24.80	11.98	3.19	12.10
Grade 10	18	23.83	11.96	24.34	12.24	-0.51	7.75
Grade 11	7	24.47	9.74	28.01	7.67	-3.54	12.72
Grade 12	4	21.60	18.75	9.45	12.37	12.15	8.21
Summary							
Program Total	277	27.21	10.66	25.54	10.53	1.67	11.37

Table 7
Fall 1989 and Fall 1988 NCE Scores
on the Metropolitan Achievement Test,
Reading Comprehension Subtest
All Students Tested
Data by School

School	Number Tested	Fall 1989 Mean	Fall 1989 S.D.	Fall 1988 Mean	Fall 1988 S.D.	Difference Mean	Difference S.D.
Target Schools							
HS-A	101	27.90	12.13	26.05	13.13	1.85	11.04
EL-C	46	37.55	12.86	38.90	11.16	-1.35	13.14
EL-D	44	25.92	9.38	27.33	8.67	-1.41	9.99
EL-G	107	28.56	12.76	27.42	12.07	1.14	12.78
EL-H	69	28.91	10.29	27.52	11.05	1.40	12.02
EL-J	24	24.96	10.98	23.60	12.45	1.36	11.25
EL-K	0	----	----	----	----	----	----
Follow the Child Service in Non-Target Schools							
EL-E	4	21.95	5.84	16.05	9.58	5.90	5.02
EL-F	0	----	----	----	----	----	----
EL-I	7	22.86	8.99	28.20	13.43	-5.34	8.84
EL-L	22	31.37	10.05	27.16	8.43	4.21	10.21
EL-M	12	34.96	9.94	26.94	13.99	8.02	13.99
EL-N	20	35.60	9.91	33.13	11.33	2.47	10.87
Parochial Schools							
HS-B	4	22.15	18.12	21.27	30.73	0.87	18.79
EL-O**	0	----	----	----	----	----	----
EL-P	34	31.96	14.07	30.55	10.22	1.41	8.51
EL-Q	13	28.07	14.20	33.28	9.57	-5.21	17.02
Program							
Total	507	29.48	12.21	28.45	12.29	1.03	11.82

**School closed 6/30/89.

Table 8
Fall 1989 and Fall 1988 NCE Scores
on the Metropolitan Achievement Test,
Reading Comprehension Subtest
Continuing Chapter 1 Students
Data by School

School	Number Tested	Fall 1989 Mean	Fall 1989 S.D.	Fall 1988 Mean	Fall 1988 S.D.	Difference Mean	Difference S.D.
Target Schools							
HS-A	52	26.11	12.48	24.35	11.87	1.75	11.14
EL-C	26	31.88	9.94	34.37	9.21	-2.48	11.44
EL-D	29	24.01	9.02	24.98	8.22	-0.98	10.46
EL-G	63	27.29	11.07	25.76	9.59	1.53	12.44
EL-H	45	28.82	7.31	24.28	9.86	4.55	10.08
EL-J	17	23.91	11.51	19.92	10.67	3.99	11.44
EL-K	0	----	----	----	----	----	----
Follow the Child Service in Non-Target Schools							
EL-E	2	17.15	2.47	8.55	2.62	8.60	5.09
EL-F	0	----	----	----	----	----	----
EL-I	1	15.40	----	29.90	----	-14.50	----
EL-L	10	27.50	7.70	26.90	6.80	0.60	8.00
EL-M	6	31.60	10.30	19.30	8.40	12.30	7.40
EL-N	10	33.05	10.37	28.03	11.65	5.02	11.77
Parochial Schools							
HS-B	1	1.00	----	1.00	----	0.00	----
EL-O**	0	----	----	----	----	----	----
EL-P	10	24.00	9.90	26.90	9.10	-3.00	6.20
EL-Q	5	29.40	14.30	30.90	10.50	-1.50	21.50
Program Total	277	27.21	10.66	25.54	10.53	1.67	11.37

**School closed 6/30/89.

TABLE 9
Spring 1990 and Spring 1989 NCE Scores
on the Metropolitan Achievement Test,
Reading Comprehension Subtest
Data by School

School	Number Tested	Spring 1990 Mean	Spring 1989 Mean	Difference Mean
Target Schools				
HS-A	54	43.20	28.30	3.90
EL-C	32	40.40	39.80	0.60
EL-D	32	35.10	30.80	4.30
EL-G	79	31.80	31.00	0.80
EL-H	43	35.70	31.70	4.00
EL-J	20	29.10	25.50	3.60
EL-K	15	39.20	37.80	1.40
Follow the Child Service in Non-Target Schools				
EL-E	4	32.90	9.50	23.50
EL-I	9	44.90	28.00	16.90
EL-L	10	35.50	31.10	4.40
EL-M	8	27.50	32.80	-5.40
EL-N	13	36.80	33.70	3.20
Parochial Schools				
HS-B	1	24.20	13.10	11.10
EL-O	0	----	----	----
EL-P	9	32.40	30.70	1.70
EL-Q	4	32.20	33.60	-1.40
Summary				
Program Total	333	34.30	31.30	3.00

FIGURE 1

End of Year
Reading Profile Summary Form
Cambridge Public Schools, Chapter 1 Program

Name _____ Gr ____ School _____

SID _____ Chapter 1 Teacher _____

I	**Reading Comprehension & Composing**				
•	Response to literature: Story frames	NA	1	2	3
•	Response to expository text: Graphic organizers	NA	1	2	3
•	Writing	NA	1	2	3
II	**Affective Dimension**				
•	Attitude toward books	NA	1	2	3
•	Voluntary reading	NA	1	2	3
•	Metacognitive knowledge of one's own reading	NA	1	2	3

MAT Spring 1990 Level _____; Percentile _____; NCE change _____;

Comments from the teacher conference:

KEY: NA = not applicable; 1 = base line; 2 = evidence of some improvement during the year; 3 = average classroom level;